基于极限学习机的极移中长期预报*

赵丹宁，雷雨
(1.宝鸡文理学院 电子电气工程学院，陕西 宝鸡 721016
2.西安邮电大学 计算机学院，陕西 西安 710121)

摘 要：传统的极移预报多是基于最小二乘外推和自回归等线性模型，但极移包含了复杂的非线性成分，线性模型的预报效果往往不甚理想。将一种新型神经网络—极限学习机（extreme learning machine，ELM） 用于极移中长期预报。首先利用最小二乘外推模型对极移序列进行拟合，获得趋势项外推值，然后采用ELM对最小二乘拟合残差进行预报，最终的极移预报值为趋势项外推值与残差预报值之和。将ELM的预报结果同反向传播（back propagation，BP）神经网络与地球定向参数预报比较活动（Earth Orientation Parameters Prediction Comparison Campaign，EOP PCC）的预报结果进行对比，结果表明：ELM用于极移中长期预报是高效可行的。

关键词：极移；预报；极限学习机；神经网络；预报模型
中图分类号：P227　　**文献标识码：**A　　**文章编号：**

0 引 言

地球极移表示地球瞬时自转轴相对于地球表面或者历元平极的运行，是表征地球自转变化的一个重要参量，简称极移动。极移的两个分量x_p、y_p、UT1-UTC与岁差、章动统称为地球定向参数（Earth orientation parameters，EOP），EOP是天球参考系和地球参考系相互转换的必需参数。甚长基线干涉（Very Long Baseline Interferometry，VLBI）和全球导航卫星系统（Global Navigation Satellite Systems，GNSS）等空间大地测量技术的极移测定精度可达0.05mas[1]。但由于复杂的资料处理过程，由空间大地测量技术获取的极移结果存在一定的延迟，而深空探测和卫星精密定轨等领域对极移的实时测量值具有重要需求，所以对极移进行短期预报就显得尤为必要，此外，极移中长期预报值在卫星自主导航和天文地球动力学研究中亦有重要应用，故极移的中长期预报也具有重要的科学及现实意义。

国内外学者提出了多种极移预报方法，如最小二乘外推[2]、自协方差模型[3]、模糊推理系统[4]、神经网络[5]、自回归模型[5-6]以及这些模型的不同组合。受大气和海洋等多种激发源的影响，极移呈现复杂的非线性、非平稳变化，因此采用神经网络等非线性方法对其进行预报在理论上更为合适，许多学者将神经网络应用于极移预报，取得了较好的预报效果[7-8]。国际EOP比较活动（EOP Prediction Comparison Campaign，EOP PCC）的结果也表明，神经网络是当前极移预报精度最高的方法之一[9]。

反向传播（back propagation，BP）神经网络是当前应用最为广泛的一种网络学习算法，在时间序列预测、图像处理和自动控制等领域均有应用，但BP神经网络存在易陷入局部极小值、学习速度慢和泛化能力弱等缺点。极限学习机（extreme learning machine，ELM）是近年发展起来的一种新型单隐含层前馈神经网络，其特点是隐含层神经元的权重可以随机给定，且不需要更新，网络学习过程仅计算输出权重，故在学习速度与泛化能力方面较传统BP神经网络具有一定优势[10-11]。正是基于此，本文将ELM神经网络应用于极移中长期预报，根据极移时间序列特征，建立了基于ELM神经网络的极移预报模型，试验结果验证了ELM神经网络应用于极移预报的有效性和实用性。

1 极限学习机原理

给定N个训练样本$(\boldsymbol{x}_i,\boldsymbol{y}_i)$，其中输入向量$\boldsymbol{x}_i=[x_{i1},x_{i2},\cdots,x_{iu}]^{\mathrm{T}}\in\mathrm{R}^u$，输出向量$\boldsymbol{y}_i=[y_{i1},y_{i2},\cdots,y_{iv}]^{\mathrm{T}}\in\mathrm{R}^v$，$u$、$v$分别为输入向量维数与输出向量维数，则具有$h$个隐含层

收稿日期：2019-**-**；　修订日期：2019-**-**
*基金项目：国家自然科学基金资助项目(11503031)
作者简介：赵丹宁(1985-)，女，陕西宝鸡人，博士，讲师，主要从事卫星自主导航方面的研究工作。E-mail:zhaodanning31@163.com

节点的单隐含层前馈神经网络能够以零误差来逼近这 N 个样本，即有下式成立[10]：

$$\boldsymbol{y}_i = \sum_{j=1}^{h} \boldsymbol{\beta}_j g(\boldsymbol{w}_j \cdot \boldsymbol{x}_i + b_j)\text{，} i = 1, 2, \cdots, N \tag{1}$$

式中， $\boldsymbol{w}_j = [w_{j1}, w_{j2}, \cdots, w_{ju}]^{\mathrm{T}}$ 为第 j 个隐含层神经元与输入层神经元间的连接权重，$\boldsymbol{\beta}_j = [\beta_{j1}, \beta_{j2}, \cdots, \beta_{jv}]^{\mathrm{T}}$ 为第 j 个隐含层神经元与输出层神经元间的连接权重，b_j 为第 j 个隐含层神经元的偏置， $g(\boldsymbol{x})$ 为激活函数。式（1）可以表示为矩阵相乘的形式：

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{\beta} \tag{2}$$

其中， $\boldsymbol{H}$ 为隐含层输出矩阵：

$$\boldsymbol{H} = \begin{bmatrix} g(\boldsymbol{w}_1 \cdot \boldsymbol{x}_1 + b_1) & \cdots & g(\boldsymbol{w}_h \cdot \boldsymbol{x}_1 + b_h) \\ \vdots & \cdots & \vdots \\ g(\boldsymbol{w}_1 \cdot \boldsymbol{x}_N + b_1) & \cdots & g(\boldsymbol{w}_h \cdot \boldsymbol{x}_N + b_h) \end{bmatrix}_{N \times h} \tag{3}$$

$\boldsymbol{\beta}$ 和 $\boldsymbol{Y}$ 为

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1^{\mathrm{T}} \\ \boldsymbol{\beta}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\beta}_N^{\mathrm{T}} \end{bmatrix}_{N \times v}\text{，} \boldsymbol{Y} = \begin{bmatrix} \boldsymbol{y}_1^{\mathrm{T}} \\ \boldsymbol{y}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{y}_N^{\mathrm{T}} \end{bmatrix}_{N \times v} \tag{4}$$

ELM神经网络的训练步骤如下[10]：1）随机选取输入层权重 $\boldsymbol{w}_j$ 、偏置 b_j ， $j = 1, 2, \cdots, h$ ；2）计算隐含层输出矩阵 $\boldsymbol{H}$ ；3）计算输出层权值 $\boldsymbol{\beta} = \boldsymbol{H}^{\dagger}\boldsymbol{Y}$ ，其中 $\boldsymbol{H}^{\dagger}$ 为矩阵 $\boldsymbol{H}$ 的Moore-Penrose广义 逆。

2 预报方法

2.1 最小二乘外推

最小二乘外推模型包含3部分：线性趋势项、钱德勒项和周年项，其数学模型为

$$f = a + \gamma t + \sum_{i=1}^{2}[c_i \cos(2\pi t / T_i) + d_i \sin(2\pi t / T_i)] \tag{5}$$

式中，T_1、T_2 分别为钱德勒项和周年项的振荡周期，取 $T_1 = 432.08\mathrm{d}$、$T_2 = 365.24\mathrm{d}$，a 、γ 为趋势项参数，c_1、d_1 为钱德勒项参数，c_2、d_2 代表周年项参数，这些未知参数可以根据最小二乘法求解。

2.2 神经网络模型构建

最小二乘拟合残差含有剩余的短周期成分，也含有各种准周期和非线性成分，对于这些成分，利用神经网络的非线性拟合能力对其进行建模与预报。依据神经网络模型构建与极移预报的实际情况，采用如下方式确定网络的输入和输出：

$$[\boldsymbol{\varepsilon}(t-k-u+1),\boldsymbol{\varepsilon}(t-k-u+2),\cdots,\boldsymbol{\varepsilon}(t-k)]\to\boldsymbol{\varepsilon}(t) \tag{6}$$

其中，$\boldsymbol{\varepsilon}(t)$ 为最小二乘拟合残差，$t=1,2,\cdots,n$，k 为预报跨度。为了提高数据利用率，采用滑动窗口方式构建训练样本：

$$\begin{cases}[\boldsymbol{\varepsilon}(1),\boldsymbol{\varepsilon}(2),\cdots,\boldsymbol{\varepsilon}(u)]\to\boldsymbol{\varepsilon}(u+k)\\ [\boldsymbol{\varepsilon}(2),\boldsymbol{\varepsilon}(3),\cdots,\boldsymbol{\varepsilon}(u+1)]\to\boldsymbol{\varepsilon}(u+k+1)\\ \qquad\vdots\\ [\boldsymbol{\varepsilon}(n-k-u+1),\boldsymbol{\varepsilon}(n-k-u+2),\cdots,\boldsymbol{\varepsilon}(n-k)]\to\boldsymbol{\varepsilon}(n)\end{cases} \tag{7}$$

相应地，在预报阶段网络的输入和输出为

$$[\boldsymbol{\varepsilon}(n-u+1),\boldsymbol{\varepsilon}(n-u+2),\cdots,\boldsymbol{\varepsilon}(n)]\to\hat{\boldsymbol{\varepsilon}}(n+k) \tag{8}$$

其中，$\hat{\boldsymbol{\varepsilon}}(n+k)$ 为神经网络模型的第 k 步预报值。

由于本文利用ELM神经网络实现从输入到输出的非线性映射，故需要事先确定激活函数类型和网络拓扑结构（输入层神经元数和隐含层神经元数），其中，激活函数选用Sigmoid函数，其表达式为

$$g(\boldsymbol{x})=1/(1+e^{-x}) \tag{9}$$

由于网络拓扑结构对网络性能起着至关重要的影响，为了确定较优的网络拓扑结构，本文通过观察网络训练误差随输入层神经元数和隐含层神经元数的变化情况来优化网络拓扑结构，即网络训练误差最小时所对应的输入层神经元数和隐含层神经元数为最优输入层和隐含层神经元数。以预报跨度30d为例，图1给出了网络训练均方根误差（root mean square error，RMSE）随输入层和隐含层神经元数的变化情况，从图1可以看到，当输入层神经元数为17、隐含层神经元数为47时，网络训练均方根误差最小。

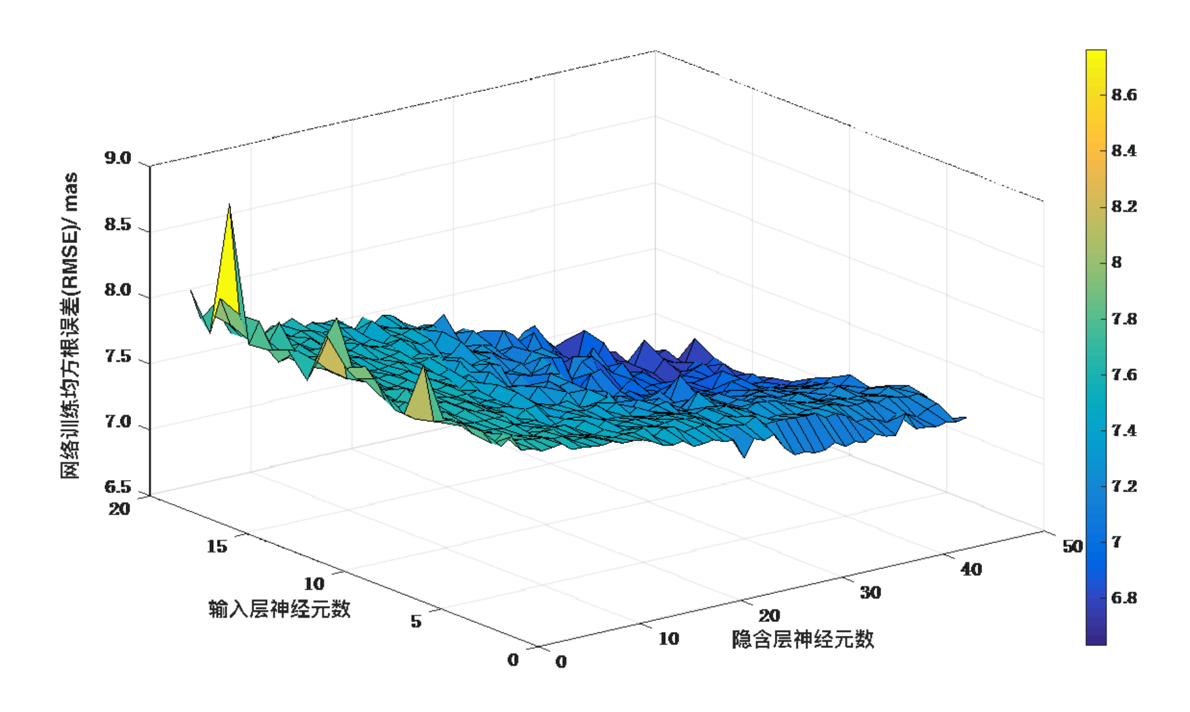


图1 网络训练均方根误差（RMSE）随输入层与输出层神经元数的变化情况

Fig.1 Changes of the network training RMSE with the neutron number of input and hidden layers

3 方法验证

为了评估ELM神经网络用于极移长期预报的性能，首先将本文的预报结果同Liao等的BP神经网络预报结果进行比较和分析[7]，为了较为客观地对二者的结果进行对比，本文预报过程中采用与Liao等相同的预报策略和误差评价指标，其中，试验数据采用IERS发布的EOP 05C04序列，基础数据长度为3yr，预报时段为2001年4月6日（MJD 52005）至2010年3月25日（MJD 55280），每隔91d预报一次，一共进行了37期的预报，预报误差评价指标选用平均绝对误差（mean absolute error，MAE），其计算公式为

$$\mathrm{MAE}_k = \frac{1}{l}\sum_{j=1}^{l}\left|P_j^k - O_j^k\right| \tag{10}$$

式中，O_j、P_j 分别表示 j 点的观测值与预测值，l 为预报期数。

图2给出了本文极移预报结果的误差曲线和Liao等预报结果的误差曲线，表1统计了二者1d、5d、10d、20d、30d、60d、90d、120d、180d、240d、300d和360d预报跨度的平均绝对误差。从图2和表1可以看出，对于 x_{p} 分量中长期预报，除了120d、180d和240d的ELM神经网络方法的预报精度稍逊于Liao等的BP神经网络方法，其余跨度的预报精度均优于Liao等的BP神经网络方法；对于 x_{p} 分量中长期预报，除了120d、180d和240d的ELM神经网络方法的预报精度稍逊于Liao等的BP神经网络方法，其余跨度的预报精度优于Liao等的BP神经网络方法或与之相当；对于 y_{p} 分量中长期预报，除了120d的预报精度稍逊于Liao等的BP神经网络方法，其余跨度的预报精度优于Liao等的BP神经网络方法或与之相当。从图2还可以发现，ELM神经网络预报误差曲线变化比较平缓，预报误差呈现逐渐增大趋势，而BP神经网络预报误差曲线跳跃较为明显，预报误差不稳定。

表1 ELM神经网络极移预报与Liao等预报结果的平均绝对误差统计（单位：mas）

Table.1 MAE statistics of the pole coordinate predictions obtained by the ELM and Liao (unit: mas)

跨度/d	1	5	10	20	30	60	90	120	180	240	300	360
Liao[15]												
x_{p}	2.12	17.64	21.84	21.21	21.70	24.83	25.96	23.99	22.74	25.40	27.70	32.59
y_{p}	1.52	11.86	16.87	17.71	18.60	21.58	23.93	23.68	27.67	33.50	31.44	31.99
ELM												
x_{p}	0.70	1.95	3.11	6.19	9.36	19.10	25.11	25.57	26.46	27.48	27.75	30.92
y_{p}	0.43	1.19	2.73	5.49	8.27	17.26	23.37	26.80	27.44	31.25	30.58	31.48

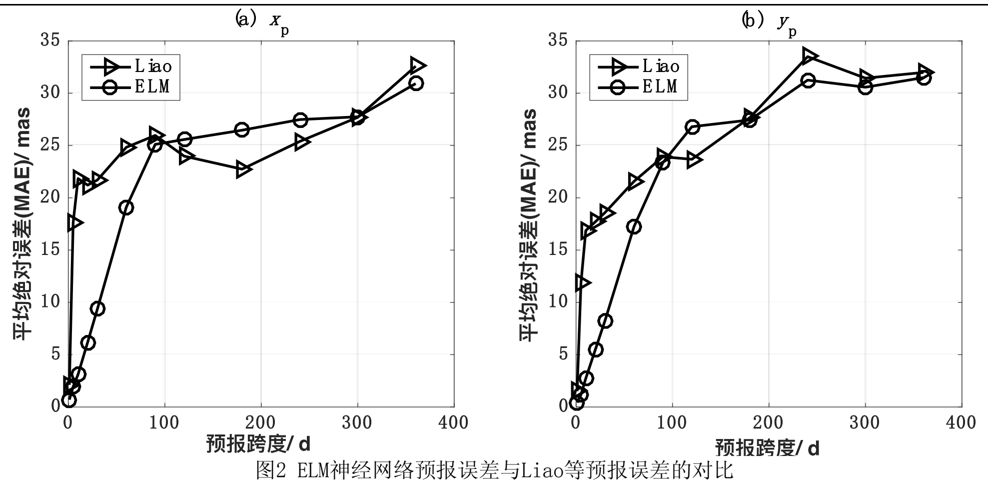


图2 ELM神经网络预报误差与Liao等预报误差的对比

Fig.2 Comparison of the prediction errors for pole coordinates obtained by the ELM and Liao

为了进一步分析ELM神经网络方法的预报性能，本文也预报了2005年10月1日至2008年2月28

日（与EOP PCC时段相同）不同预报跨度的极移值，并将本文预报结果的平均绝对误差与EOP PCC预报结果的平均绝对误差进行对比分析和分析[9]，如图3～图5所示，图中数字代表参与EOP PCC的不同团队的编号，不同颜色曲线表示不同团队预报结果的平均绝对误差，参与EOP PCC的团队详细情况详见文献[9]，其中，极移超短期（1～10d）预报精度较高的团队是Kalarus（编号061，黑色实线）和IERS EOP产品中心（编号071，绿色实线）；短期（1～30d）预报精度较高的团队是Kalarus（编号061，黑色实线）、Zotov（编号091，浅蓝色实线）和Kosek（编号051，灰蓝色实线）；中长期（1～500d）预报精度较高的团队是Kalarus（编号061，黑色实线）和Kosek（编号051，灰蓝色实线）。

从图3～图5可以看到，对于1～10d的超短期预报，ELM神经网络方法的预报精度仅低于排在第一位的Kalarus团队的预报精度，优于其他团队的预报精度；对于1～30d的短期预报，ELM神经网络方法的预报精度大体处于所有团队的居中位置；对于1～500d的中长期预报，ELM神经网络方法的预报精度低于排在第一位的Kalarus团队的预报精度，稍逊于排在第二位的Kosek团队的预报精度或与之相当，优于其他团队的预报精度。

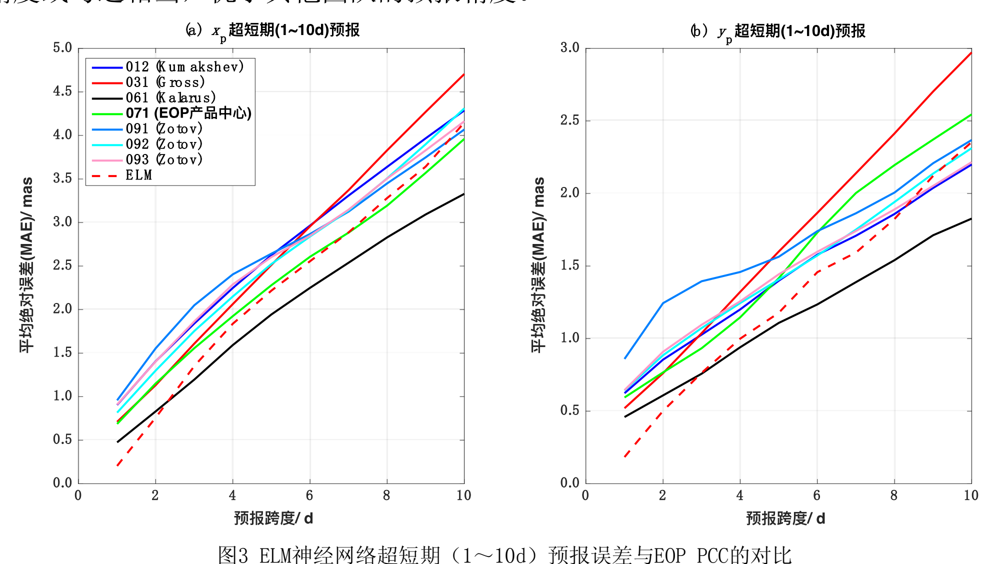


图3 ELM神经网络超短期（1～10d）预报误差与EOP PCC的对比

Fig.3 Comparison of the ultra short-term (1～10d) prediction errors obtained by the ELM and EOP PCC

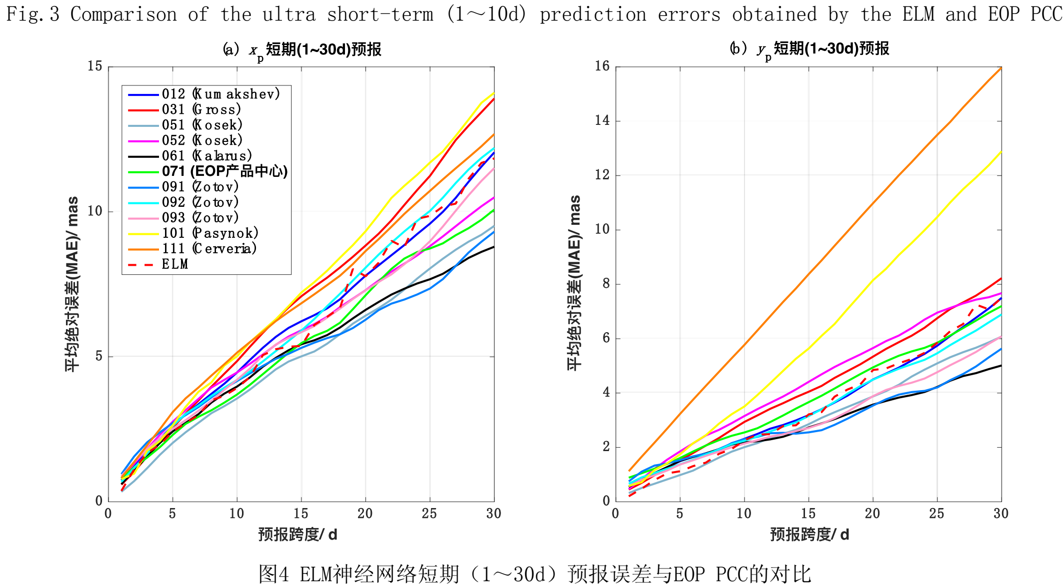


图4 ELM神经网络短期（1～30d）预报误差与EOP PCC的对比

Fig.4 Comparison of the short-term (1～30d) prediction errors obtained by the ELM and EOP PCC

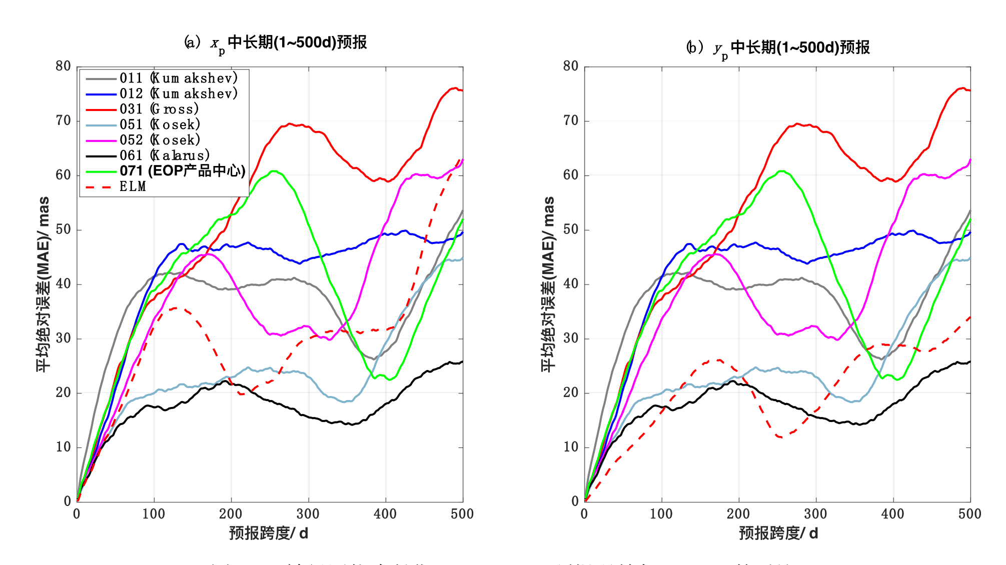


图5 ELM神经网络中长期（1～500d）预报误差与EOP PCC的对比

Fig.5 Comparison of the medium- and long-term (1～500d) prediction errors obtained by the ELM and EOP PCC

4 结 论

本文针对极移序列非线性、非平稳变化的特性，将一种新型ELM神经网络用于极移中长期预报。通过预报试验可以发现，ELM神经网络用于极移中长期预报可以获得较好的预报效果。与BP神经网络预报结果的对比表明，ELM神经网络的预报精度优于BP神经网络的预报精度；与EOP PCC预报结果的比较发现，ELM神经网络的中长期预报精度优于大部分参与EOP PCC团队的预报精度。

参考文献

1 Dill R, Dobslaw H. Short-term polar motion forecasts from earth system modeling data [J]. Journal of Geodesy, 2010, 84(9): 529-536.

2 Guo J Y, Li Y B, Dai C L, et al. A technique to improve the accuracy of Earth orientation prediction algorithms based on least squares extrapolation [J]. Journal of Geodynamics, 2013, 70(10): 36-48.

3 Kosek W, McCarthy D D, Luzum B J. Possible improvement of Earth orientation forecast using autocovariance prediction procedures [J]. Journal of Geodesy, 1998, 72(4): 189-199.

4 Akyilmaz O, Kutterer H. Prediction of Earth rotation parameters by fuzzy inference systems [J]. Journal of Geodesy, 2004, 78(1-2): 82-193.

5 Schuh H, Ulrich M, Egger D, et al. Prediction of Earth orientation parameters by artificial neural networks [J]. Journal of Geodesy, 2002, 76(5): 247-258.

6 许雪晴，周永宏.地球定向参数高精度预报方法研究[J].飞行器测控学报，2010，29（2）：70-76.
Xu Xueqing, Zhou Yonghong. High precision prediction method of Earth orientation parameters [J]. Journal of Spacecraft TT & Technology, 2010, 29(2): 70-76.

7 Liao D C, Wang Q J, Zhou Y H, et al. Long-term prediction of the Earth orientation parameters by the artificial neural network technique[J]. Journal of Geodynamics, 2012, 62(8): 87-925.

8 赵丹宁，雷雨，蔡宏兵. 利用经验模态分解提高极移预报精度[J].天文研究与技术，2018，15（2）：140-150.
Zhao Danning, Lei Yu, Cai Hongbing,. Enhancement of the prediction accuracy of pole coordinates with empirical mode decomposition [J]. ASTRONOMICAL RESEARCH & TECHNOLOGY, 2018, 15(2): 140-150.

9 Kalarus M, Schuh H, Kosek W, et al. Achievements of the Earth orientation parameters prediction comparison campaign[J]. Journal of Geodesy, 2010, 84(10): 587-596.

10 Huang G B, Zhou Q Y, Siew C K. Extreme learning machine: theory and applications [J]. Neurocomputing, 2006, 70(1-3): 489-501.

11 赵丹宁，雷雨. 利用端点延拓提高 LS + NN 模型的 UT1-UTC 预报精度[J].天文研究与技术，2019，16（3）：294-299.
Zhao Danning, Lei Yu. Improving the performance of the LS+NN model for UT1-UTC forecast with the edge extension [J]. ASTRONOMICAL RESEARCH & TECHNOLOGY, 2019, 16(3): 294-299.

Medium- and long-term prediction of polar motion using extreme learning machine

Zhao Danning, Lei Yu

(1 School of Electrical & Electronic Engineering, Baoji University of Arts and Sciences, Baoji 721013, China

2 School of Computer Science & Technology, Xi'an University of Posts & Telecommunications, Xi'an 710121, China)

Abstract: Linear models such as least squares (LS) extrapolation and autoregressive (AR) models get used to prediction Earth's polar motion. However, polar motion predictions obtained by linear models are not fully satisfactory owing to the non-linear and non-stationary variations in polar motion. In this paper, the extreme learning machine (ELM), a new training algorithm for single hidden layer feedback neural network (SLFN), is employed to prediction medium- and long-term polar motion. The LS extrapolation model is first employed to separate a linear trend, Chandler and annual wobbles (CW and AW) from observed polar motion time-series. The extracted linear trend, CW and AW are then predicted by using the above-mentioned LS extrapolation model. Secondly, the ELM is used for modeling and forecasting the residual time-series obtained by subtracting the LS extrapolation curve from the original polar motion data. In ELM training some special issues such as generation of training patterns and topology of the network for enhanced medium- and long-term prediction of polar motion are taken into consideration. The subsequently forecasted residuals are then added to the LS extrapolation so as to obtain the predicted values of polar motion. In order to demonstrate the effectiveness of the presented method, the results of the predictions of polar motion are analyzed and compared with those obtained by the back propagation (BP) neural network and Earth Orientation Parameters Prediction Comparison Campaign (EOP PCC). The results show that the predictions are comparable with or better than those obtained by other existing prediction methods and techniques in terms of the mean absolute error (MAE).

Key words: Polar motion; Prediction; Extreme learning machine; Neural network; Prediction model